Linear Quadratic Problems with Indefinite Cost for Discrete Time Systems

Ran, A. C. M.; Trentelman, H. L.

Published in:
SIAM Journal on Matrix Analysis and Applications

DOI:
10.1137/0614055

SIAM J. MATRIX ANAL. APPL.
Vol. 14, No. 3, pp. 776–797, July 1993

013

LINEAR QUADRATIC PROBLEMS WITH INDEFINITE COST FOR DISCRETE TIME SYSTEMS*

A. C. M. RAN† AND H. L. TRENTELMAN‡

Abstract. This paper deals with the discrete-time, infinite-horizon linear quadratic problem with indefinite cost criterion. Given a discrete-time linear system, an indefinite cost-functional and a linear subspace of the state space, the problem of minimizing the cost-functional over all inputs that force the state trajectory to converge to the given subspace is considered. A geometric characterization of the set of all Hermitian solutions of the discrete-time algebraic Riccati equation is given. This characterization forms the discrete-time counterpart of the well-known geometric characterization of the set of all real symmetric solutions of the continuous-time algebraic Riccati equation as developed by Willems [*IEEE Trans. Automat. Control*, 16 (1971), pp. 621–634] and Coppel [*Bull. Austral. Math. Soc.*, 10 (1974), pp. 377–401]. In the set of all Hermitian solutions of the Riccati equation the solution that leads to the optimal cost for the above-mentioned linear quadratic problem is identified. Finally, necessary and sufficient conditions for the existence of optimal controls are given.

Key words. linear quadratic optimal control, indefinite cost functional, discrete-time, Riccati equation

AMS subject classifications. 93C05, 93C35, 93C60

1. Introduction. This paper has two main goals. First, we want to establish the discrete-time counterpart of the by now "classical" geometric characterization of the lattice of real symmetric solutions of the continuous-time algebraic Riccati equation as given in [1] and [8]. Subsequently, we want to apply these results to the discrete-time linear quadratic optimization problem with linear endpoint constraints. Given a discrete-time linear system, the latter problem consists of minimizing a general *indefinite* quadratic cost-functional over the class of input functions that force the state trajectory to converge to an a priori given subspace of the state space (or, equivalently, that force a given linear function of the state to converge to zero). A complete treatment of this optimization problem for the continuous-time case was given only very recently in [5] and [6].

With respect to our first goal, it will be shown that, as in the continuous-time case, if the algebraic Riccati equation has at least one Hermitian solution, then it has a smallest one and a largest one. Furthermore, any Hermitian solution of the algebraic Riccati equation can be written as a "linear combination" of these extremal solutions. In order to derive these results we will make use of the characterization of all Hermitian solutions of the discrete-time Riccati equation in terms of certain invariant Lagrangian subspaces, as established in [4] (see also [2]).

With respect to our second goal, we want to note that compared to the usual discrete-time linear quadratic optimal control problems, our problem formulation introduces generalizations into two independent directions. First, in contrast to the existing literature on this subject, we do not require the quadratic form in the cost-functional to be positive semidefinite (the "linear quadratic regulator problem"). Instead, the quadratic form is allowed to be indefinite. Second, our problem formula-

* Received by the editors December 19, 1990; accepted for publication (in revised form) November 22, 1991.

† Faculty of Mathematics and Computing Science, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands (ran@cs.vu.nl).

‡ Mathematics Institute, P.O. Box 800, 9700 AV Groningen, the Netherlands (trentelman@rug.nl).

tion includes a fixed, but arbitrary, linear endpoint constraint, in the sense that the optimization is performed over the class of all input functions that force the state trajectory to converge to an a priori given subspace. A solution to the usual zero endpoint problem (in which the optimal state trajectory is required to converge to the origin) can thus be obtained from our results by setting this subspace to be equal to the zero subspace. On the other hand, a solution to the free endpoint problem (no constraint on the optimal state trajectory) can be obtained from our results by taking the subspace to be equal to the entire state space.

The outline of this paper is as follows. In §2 we shall formulate the optimization problem that we want to consider. This section also contains a statement of the main result of this paper, that is, a characterization of the optimal cost, necessary and sufficient conditions for the existence of optimal controls, and an expression for the optimal state feedback control law. In §3 we shall establish the characterization of the set of all Hermitian solutions of the discrete-time algebraic Riccati equation as announced above. Finally, in §4 we shall give a proof of the main result as stated in §2.

2. Problem statement and main results. In this paper, we will consider the discrete-time system

$$x_{k+1} = Ax_k + Bu_k, \tag{2.1}$$

where the state variable x_k takes its values in $\mathcal{C}^n$ and the input variable u_k takes its values in $\mathcal{C}^m$. In (2.1) we have $A \in \mathcal{C}^{n\times n}$ and $B \in \mathcal{C}^{n\times m}$. As a standing assumption, we take (A, B) to be a controllable pair. We will consider optimization problems of the type

$$\inf \sum_{k=0}^{\infty} \begin{pmatrix} x_k \\ u_k \end{pmatrix}^* \begin{pmatrix} Q & C^* \\ C & R \end{pmatrix} \begin{pmatrix} x_k \\ u_k \end{pmatrix}. \tag{2.2}$$

Here, R, Q, and C are complex matrices of appropriate dimensions, and $R = R^*$, $Q = Q^*$. In this paper, a standing assumption will be that R is nonsingular. The expression (2.2) of course needs some explanation. For any $x_0 \in \mathcal{C}^n$ and any control sequence $u = \{u_k\}_{k=0}^{\infty}$, we define

$$J_T(x_0, u) := \sum_{k=0}^{T} \begin{pmatrix} x_k \\ u_k \end{pmatrix}^* \begin{pmatrix} Q & C^* \\ C & R \end{pmatrix} \begin{pmatrix} x_k \\ u_k \end{pmatrix}. \tag{2.3}$$

Let

$$U(x_0) := \left\{ u \mid \lim_{T\to\infty} J_T(x_0, u) \text{ exists in } \mathcal{R} \cup \{-\infty, +\infty\} \right\},$$

and for any control sequence $u \in U(x_0)$, define the associated cost by

$$J(x_0, u) := \lim_{T\to\infty} J_T(x_0, u). \tag{2.4}$$

The optimization problem of minimizing the cost functional (2.4) over the class of inputs $U(x_0)$ is called the *free endpoint linear quadratic problem.* The optimal cost associated with this problem is equal to

$$V_f(x_0) := \inf_{u\in U(x_0)} J(x_0, u). \tag{2.5}$$

Compare this problem with the usual *zero endpoint problem*, where instead of $U(x_0)$, the cost-functional is minimized over the class of all inputs that force the corresponding state trajectory to converge to the origin, i.e., over

$$U_s(x_0) := \left\{ u \in U(x_0) | \lim_{k \longrightarrow \infty} x(x_0, u)_k = 0 \right\}. \tag{2.6}$$

The associated optimal cost is given by

$$V_+(x_0) := \inf_{u \in U_s(x_0)} J(x_0, u).$$

In the present paper, we will study a generalization of the above two linear quadratic problems, the *linear quadratic problem with linear endpoint constraints*. Given a linear subspace $\mathcal{L}$ of the state space $\mathcal{C}^n$, the latter problem consists of minimizing the cost functional (2.4) over all inputs u that force the state trajectory to converge to the subspace $\mathcal{L}$:

$$U_{\mathcal{L}}(x_0) := \left\{ u \in U(x_0) | \lim_{k \longrightarrow \infty} d(x(x_0, u)_k, \mathcal{L}) = 0 \right\}. \tag{2.7}$$

In the above, for a given point $x \in \mathcal{C}^n$, $d(x, \mathcal{L})$ denotes the usual distance from the point x to the subspace $\mathcal{L}$. The optimal cost for the latter problem is given by

$$V_{\mathcal{L}}(x_0) := \inf_{u \in U_{\mathcal{L}}(x_0)} J(x_0, u). \tag{2.8}$$

Obviously, both the free endpoint problem and the fixed endpoint problem are special cases of the latter problem formulation: take $\mathcal{L} = \mathcal{C}^n$ and $\mathcal{L} = 0$, respectively.

An important role will be played by the set of Hermitian solutions of the discrete-time algebraic Riccati equation

$$P = A^*PA + Q - (C + B^*PA)^*(R + B^*PB)^{-1}(C + B^*PA). \tag{2.9}$$

P is called a solution of the algebraic Riccati equation if $R + B^*PB$ is nonsingular and if P satisfies (2.9). Besides controllability of (A, B) and nonsingularity of R, we shall assume throughout that $A - BR^{-1}C$ is nonsingular and that $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$, where

$$\Psi(z) := \begin{pmatrix} B^*(Iz^{-1} - A^*)^{-1} & I \end{pmatrix} \begin{pmatrix} Q & C^* \\ C & R \end{pmatrix} \begin{pmatrix} (Iz - A)^{-1}B \\ I \end{pmatrix}.$$

Here, $\mathcal{T}$ denotes the unit circle. Let us comment on these assumptions. The nonsingularity of R and $A - BR^{-1}C$ is mainly assumed for technical reasons. These assumptions allow us to give a description of the set of Hermitian solutions of (2.9) in terms of invariant subspaces of a certain matrix (see §3) rather than in terms of a pencil of matrices (compare, e.g., [12]). Note that, in contrast to most literature on the continuous-time case, we do not assume $R > 0$, but only R nonsingular. (If in the continuous-time case the condition $R \geq 0$ is violated, then we have $V_+(x_0) = -\infty$ for every initial state x_0. In the discrete-time case the condition $R \geq 0$ is *not* necessary for boundedness from below of the cost-functional $J(x_0, u)$.) The positivity of $\Psi(\eta)$ for some $\eta \in \mathcal{T}$ should be contrasted with the continuous-time case. The corresponding matrix-valued function for the continuous-time case is given by

$$\tilde{\Psi}(z) := \begin{pmatrix} B^*(Iz + A^*)^{-1} & I \end{pmatrix} \begin{pmatrix} Q & C^* \\ C & R \end{pmatrix} \begin{pmatrix} (Iz - A)^{-1}B \\ I \end{pmatrix}.$$

Taking the limit as z goes to infinity, this function is seen to have value R at infinity. So, under the usual assumption of the continuous-time case, $R > 0$, a similar assumption is automatically satisfied: there is an η on the imaginary axis such that $\tilde{\Psi}(\eta) > 0$.

Under the assumptions outlined in the previous paragraph, the set of Hermitian solutions of (2.9) will turn out to have a maximal element P_+ and a minimal element P_- (see (3.5) below). (For the existence of both a maximal and a minimal Hermitian solution, the controllability of (A, B) is a natural sufficient condition. Assuming only *stabilizability* of (A, B) would only guarantee the existence of a maximal solution.) Put $\Delta := P_+ - P_-$. Let us denote by A_+ and A_- the matrices

$$A_+ := A - B(R + B^*P_+B)^{-1}(C + B^*P_+A),$$

$$A_- := A - B(R + B^*P_-B)^{-1}(C + B^*P_-A).$$

It will be seen that $\sigma(A_+) \subset \bar{\mathcal{D}}$ and $\sigma(A_-) \subset \bar{\mathcal{D}}^e$ (here $\mathcal{D}$ denotes the open unit disk, $\mathcal{D}^e$ denotes the exterior of the closed unit disc, and for any matrix A, $\sigma(A)$ denotes the set of eigenvalues of A). Given a subspace $\mathcal{L}$ of $\mathcal{C}^n$ we introduce the subspace

$$\mathcal{V}(\mathcal{L}) := \langle \mathcal{L} \cap \ker P_- \mid A_- \rangle \cap \mathcal{X}_+(A_-). \tag{2.10}$$

(Here, for a matrix A, $\mathcal{X}_+(A)$ denotes the spectral subspace of A corresponding to its eigenvalues in $\mathcal{D}^e$; likewise, we shall denote $\mathcal{X}_0(A)$, $\mathcal{X}_-(A)$ the spectral subspaces of A corresponding to its eigenvalues in $\mathcal{T}$, respectively, $\mathcal{D}$.) As usual, the notation $\langle \mathcal{W} \mid A \rangle$, where $\mathcal{W}$ is a subspace of $\mathcal{C}^n$ and A an $n \times n$ matrix, denotes the largest A-invariant subspace in $\mathcal{W}$. If H is a matrix such that $\ker H = \mathcal{L} \cap \ker P_-$, then $\mathcal{V}(\mathcal{L})$ is the undetectable subspace of the pair (H, A_-) with respect to the stability set $\bar{\mathcal{D}}$. Now, let

$$P_{\mathcal{L}} := P_-\pi_{\mathcal{V}(\mathcal{L})} + P_+(I - \pi_{\mathcal{V}(\mathcal{L})}), \tag{2.11}$$

where $\pi_{\mathcal{V}(\mathcal{L})}$ is the projection onto $\mathcal{V}(\mathcal{L})$ along $(\Delta\mathcal{V}(\mathcal{L}))^{\perp}$. It will turn out that $P_{\mathcal{L}}$ is a solution of (2.9). Finally, if $\mathcal{L}$ is a subspace of $\mathcal{C}^n$ and if P is a Hermitian $n \times n$ matrix, then we will say that P is *negative semidefinite on* $\mathcal{L}$ if the following conditions hold:

- for all $x_0 \in \mathcal{L} : x_0^*Px_0 \leq 0$,
- for all $x_0 \in \mathcal{L} : x_0^*Px_0 = 0 \Leftrightarrow Px_0 = 0$.

According to this definition, a Hermitian matrix P is negative semidefinite on $\mathcal{C}^n$ if and only if it is negative semidefinite in the usual sense. Furthermore, any Hermitian matrix is negative semidefinite on the zero subspace $\{0\}$. The main result of this paper can now be formulated as follows.

THEOREM 2.1. *Suppose (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Assume further that* (2.9) *has at least one Hermitian solution and assume P_- is negative semidefinite on $\mathcal{L}$. Then we have*

(i) *$V_{\mathcal{L}}(x_0)$ is finite for all x_0, and $V_{\mathcal{L}}(x_0) = x_0^*P_{\mathcal{L}}x_0$,*

(ii) *for all x_0 there is an input u^+ such that $V_{\mathcal{L}}(x_0) = J(x_0, u^+)$ if and only if* $\ker \Delta \subseteq \mathcal{L} \cap \ker P_-$; *in that case u^+ is unique and is given by the state feedback control law*

$$u_k^+ = -(R + B^*P_{\mathcal{L}}B)^{-1}(C + B^*P_{\mathcal{L}}A)x_k.$$

This result is the discrete-time analogue of [6, Thm. 4.1]. We stress that the above theorem also provides information on the free endpoint problem and on the zero endpoint problem. Indeed, for the free endpoint problem we set $\mathcal{L} = \mathcal{C}^n$. The corresponding subspace $\mathcal{V}(\mathcal{L})$ is then equal to $\mathcal{V} = \langle \ker P_- \mid A_- \rangle \cap \mathcal{X}_+(A_-)$. Define

$$P_f := P_- \pi_{\mathcal{V}} + P_+(I - \pi_{\mathcal{V}}),$$

where, again, $\pi_{\mathcal{V}}$ is the projection onto $\mathcal{V}$ along $(\Delta\mathcal{V})^\perp$. P_f is a solution of (2.9) and we find the following corollary.

COROLLARY 2.2. *Suppose (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Assume further that* (2.9) *has at least one Hermitian solution and assume that $P_- \leq 0$. Then we have*

(i) *$V_f(x_0)$ is finite for all x_0, and $V_f(x_0) = x_0^* P_f x_0$,*

(ii) *for all x_0 there is an input u^+ such that $V_f(x_0) = J(x_0, u^+)$ if and only if* $\ker \Delta \subseteq \ker P_-$; *in that case u^+ is unique and is given by the state feedback control law*

$$u_k^+ = -(R + B^* P_f B)^{-1}(C + B^* P_f A)x_k.$$

The above corollary is the discrete-time analogue of [5, Thm. 5.1]. In order to get the corresponding result on the zero endpoint problem, we set $\mathcal{L} = \{0\}$. The corresponding subspace $\mathcal{V}(\mathcal{L})$ is then equal to $\mathcal{V} = \{0\}$ and we find that the relevant solution of (2.9) is equal to P_+. Thus we find the following discrete-time version of [8, Thm. 7].

COROLLARY 2.3. *Suppose (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Assume further that* (2.9) *has at least one Hermitian solution. Then we have*

(i) *$V_+(x_0)$ is finite for all x_0, and $V_+(x_0) = x_0^* P_+ x_0$,*

(ii) *for all x_0 there is an input u_+ such that $V_+(x_0) = J(x_0, u_+)$ if and only if* $\Delta > 0$; *in that case u_+ is unique and is given by the state feedback control law*

$$u_k^+ = -(R + B^* P_+ B)^{-1}(C + B^* P_+ A)x_k.$$

In this paper, we shall give a proof of Theorem 2.1. The proof that we shall give basically follows the line of [6]; the details will be provided in §4. In §3 we give a description of all solutions of (2.9) in terms of P_- and P_+. The continuous-time analogue of this description is due to Coppel [1]. The argument here is somewhat more complicated and uses ideas from [2] and [4].

3. Description of solutions of the algebraic Riccati equation. Consider the discrete-time algebraic Riccati equation (2.9). In addition to the controllability of (A, B) and the assumption that R is nonsingular, we assume throughout that $A - BR^{-1}C$ is nonsingular and $\Psi(\eta) > 0$ for some η on the unit circle. We are then in a position to apply [4, Thm. 4.1] (see also [2, Thm. 4.4]), which gives a description of solutions of the algebraic Riccati equation in terms of certain invariant subspaces. To be precise, put

(3.1)

$$T = \begin{pmatrix} A - BR^{-1}C + BR^{-1}B^*(A - BR^{-1}C)^{*-1}(Q - C^*R^{-1}C) & -BR^{-1}B^*(A - BR^{-1}C)^{*-1} \\ -(A - BR^{-1}C)^{*-1}(Q - C^*R^{-1}C) & (A - BR^{-1}C)^{*-1} \end{pmatrix}.$$

The next theorem provides necessary and sufficient conditions for the existence of a Hermitian solution of (2.9), and two different descriptions of the set of all Hermitian solutions. Both descriptions are in terms of certain invariant subspaces of T.

THEOREM 3.1. *Assume (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Then the following statements are equivalent:*

(i) *There exists a Hermitian solution of* (2.9),

(ii) *T has an invariant subspace $\mathcal{M}$ such that*

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \mathcal{M} = \mathcal{M}^{\perp}, \tag{3.2}$$

(iii) *the partial multiplicities of T (i.e., the sizes of the Jordan blocks in the Jordan form of T) corresponding to its eigenvalues on the unit circle $\mathcal{T}$ are all even,*

(iv) *$\Psi(z) \geq 0$ for all z on the unit circle $\mathcal{T}$.*

DESCRIPTION 1. *Assume that one of the above equivalent conditions hold. Then any T-invariant subspace $\mathcal{M}$ for which* (3.2) *holds is of the form*

$$\mathcal{M} = \operatorname{im} \begin{pmatrix} I \\ P \end{pmatrix} \tag{3.3}$$

for some Hermitian solution P of (2.9), *and, conversely, if $P = P^*$ solves* (2.9), *then $\mathcal{M}$ given by* (3.3) *is T-invariant and satisfies* (3.2).

DESCRIPTION 2. *Assume that one of the above equivalent conditions hold. Then for every T-invariant subspace $\mathcal{N}$ with the property that $\sigma(T \mid \mathcal{N}) \subset \mathcal{D}^e$ there is a unique solution $P = P^*$ of* (2.9) *with*

$$\operatorname{im} \begin{pmatrix} I \\ P \end{pmatrix} \cap \mathcal{X}_+(T) = \mathcal{N}. \tag{3.4}$$

Conversely, for every Hermitian solution P of (2.9) *the subspace $\mathcal{N}$ given by* (3.4) *is T-invariant and has the property that $\sigma(T \mid \mathcal{N}) \subset \mathcal{D}^e$. Here $\mathcal{X}_+(T)$ denotes the sum of the generalized eigenspaces of T with respect to its eigenvalues in $\mathcal{D}^e$.*

Now let P be any Hermitian solution of (2.9). Then it is a straightforward calculation to see that

$$\Psi(z) = \Xi(\bar{z}^{-1})^*(R + B^*PB)\Xi(z),$$

where

$$\Xi(z) = I + (R + B^*PB)^{-1}(C + B^*PA)(Iz - A)^{-1}B$$

(see, e.g., [2]). Consequently, for any Hermitian solution P of (2.9) we have $R + B^*PB \geq 0$. Combined with the fact that $R + B^*PB$ is also nonsingular, we find that if (i)–(iv) in Theorem 3.1 hold then we have $R + B^*PB > 0$ for *any* solution $P = P^*$ of (2.9) (see also [2, Thm. 2.5]).

Let P_+ and P_- be the unique solutions for which

$$\operatorname{im} \begin{pmatrix} I \\ P_+ \end{pmatrix} \cap \mathcal{X}_+(T) = \{0\}, \qquad \operatorname{im} \begin{pmatrix} I \\ P_- \end{pmatrix} \cap \mathcal{X}_+(T) = \mathcal{X}_+(T), \tag{3.5}$$

respectively. We shall show that P_+ is the maximal solution and P_- the minimal solution of the equation (2.9). First we prove a lemma.

LEMMA 3.1. *Let P_- be the solution defined by* (3.5) *and suppose P is an arbitrary Hermitian solution. Introduce*

$$\begin{aligned} A_- &= A - B(R + B^*P_-B)^{-1}(C + B^*P_-A), \\ S_- &= R + B^*P_-B. \end{aligned}$$

Then $X := P - P_-$ satisfies the algebraic Riccati equation

$$X = A_-^* X A_- - A_-^* X B(S_- + B^*XB)^{-1}B^*XA_-. \tag{3.6}$$

Conversely, any Hermitian solution X of (3.6) *gives a solution of* (2.9) *via $P = X + P_-$.*

Proof. Introduce the following matrices:

$$\begin{aligned} S_- &:= R + B^*P_-B, & S &:= R + B^*PB, \\ E_- &:= C + B^*P_-A, & E &:= C + B^*PA, \\ L_- &:= S_-^{-1}E_-, & L &:= S^{-1}E. \end{aligned}$$

To prove (3.6), compute

$$\begin{aligned} X - A_-^*XA_- &= X - (A - BL_-)^*X(A - BL_-) \\ &= (P - P_-) - A^*(P - P_-)A + A^*(P - P_-)BL_- \\ &\quad + L_-^*B^*(P - P_-)A - L_-^*B^*(P - P_-)BL_-. \end{aligned} \tag{3.7}$$

Since P and P_- solve (2.9), we have

$$X - A^*XA = E_-^*S_-^{-1}E_- - E^*S^{-1}E. \tag{3.8}$$

Furthermore,

$$\begin{aligned} A^*(P - P_-)BL_- &= (A^*PB - A^*P_-B)L_- \\ &= (E^* - E_-^*)L_- \\ &= (E^* - E_-^*)S_-^{-1}E_- \end{aligned}$$

and

$$L_-^*B^*(P - P_-)BL_- = E_-^*S_-^{-1}B^*(P - P_-)BS_-^{-1}E_-.$$

Using these equalities and (3.8) in (3.7), we obtain

$$\begin{aligned} X - A_-^*XA_- &= E_-^*S_-^{-1}E_- - E^*S^{-1}E + (E^* - E_-^*)S_-^{-1}E_- \\ &\quad + E_-^*S_-^{-1}(E - E_-) - E_-^*S_-^{-1}B^*(P - P_-)BS_-^{-1}E_- \\ &= -E_-^*S_-^{-1}E_- - E^*S^{-1}E + E^*S_-^{-1}E_- + E_-^*S_-^{-1}E \\ &\quad - E_-^*S_-^{-1}B^*(P - P_-)BS_-^{-1}E_- \\ &= -E_-^*S_-^{-1}(S_- + B^*(P - P_-)B)S_-^{-1}E_- - E^*S^{-1}E \\ &\quad + E^*S_-^{-1}E_- + E_-^*S_-^{-1}E \\ &= -E_-^*S_-^{-1}SS_-^{-1}E_- - E^*S^{-1}E + E^*S_-^{-1}E_- + E_-^*S_-^{-1}E \\ &= -(E^* - E_-^*S_-^{-1}S)S^{-1}(E - SS_-^{-1}E_-). \end{aligned} \tag{3.9}$$

Here we used the fact that $S_- + B^*XB = S$. Moreover,

$$SS_-^{-1} = (R + B^*PB)(R + B^*P_-B)^{-1} = I + B^*XBS_-^{-1},$$

so

$$\begin{aligned} E - SS_-^{-1}E_- &= E - E_- - B^*XBS_-^{-1}E_- \\ &= B^*XA - B^*XBS_-^{-1}E_- \\ &= B^*X(A - BS_-^{-1}E_-) \\ &= B^*XA_-. \end{aligned}$$

Hence, from (3.9), we see

$$X - A_-^*XA_- = -A_-^*XB(S_- + B^*XB)^{-1}B^*XA_-,$$

which proves (3.6). The converse follows by a similar computation. □

With a similar argument one shows that the following lemma is true.

LEMMA 3.2. *Let P_+ be the solution introduced above; suppose P is an arbitrary Hermitian solution. Introduce*

$$\begin{aligned} A_+ &= A - B(R + B^*P_+B)^{-1}(C + B^*P_+A), \\ S_+ &= R + B^*P_+B. \end{aligned}$$

Then $X = P - P_+$ satisfies the algebraic Riccati equation

$$X = A_+^*XA_+ - A_+^*XB(S_+ + B^*XB)^{-1}B^*XA_+. \tag{3.10}$$

Conversely, any solution of (3.10) *gives a solution of* (2.9) *via $P = X + P_+$.*

According to Theorem 3.1, the subspaces

$$\operatorname{im}\begin{pmatrix} I \\ P_+ \end{pmatrix} \quad \text{and} \quad \operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix}$$

are T-invariant. The following lemma states that the restrictions of T to these subspaces are semistable and semi-antistable, respectively.

LEMMA 3.3.

$$\sigma\left(T \mid \operatorname{im}\begin{pmatrix} I \\ P_+ \end{pmatrix}\right) \subset \bar{\mathcal{D}}, \qquad \sigma\left(T \mid \operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix}\right) \subset \bar{\mathcal{D}}^e.$$

Proof. The first statement is an immediate consequence of (3.5). To prove the second statement, define

$$J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Note that T is J-unitary, i.e., $T^*JT = J$. This implies that T is invertible and that $T^{-1} = -JT^*J$. Now assume that

$$x \in \operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix}, \quad Tx = \lambda x \quad \text{and} \quad \lambda \in \mathcal{D}.$$

Then $T^{-1}x = \lambda^{-1}x$, from which $T^*Jx = \lambda^{-1}Jx$. This implies that $Jx \in \mathcal{X}_+(T^*)$, since $\lambda^{-1} \in \mathcal{D}^e$. We also have

$$Jx \in J\operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix} = \left(\operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix}\right)^\perp.$$

Using (3.5) the latter implies $Jx \in \mathcal{X}_+(T)^\perp$. Since $\mathcal{X}_+(T)^\perp = \mathcal{X}_0(T^*) \oplus \mathcal{X}_-(T^*)$ we obtain $Jx = 0$, so $x = 0$. □

It is a straightforward but tedious calculation to show that

$$T \mid \operatorname{im}\begin{pmatrix} I \\ P_+ \end{pmatrix}, \qquad T \mid \operatorname{im}\begin{pmatrix} I \\ P_- \end{pmatrix}$$

are similar to A_+ and A_-, respectively (see, e.g., [10] and [11]). Consequently, $\sigma(A_+) \subset \bar{\mathcal{D}}$ and $\sigma(A_-) \subset \bar{\mathcal{D}}^e$. Introduce

$$A_P = A - B(R + B^*PB)^{-1}(C + B^*PA). \tag{3.11}$$

The following theorem states that P_- is the smallest Hermitian solution of (2.9). Likewise, P_+ is the largest Hermitian solution of (2.9). Furthermore, $P = P_-$ is the only Hermitian solution with the property that $\sigma(A_P) \subset \bar{\mathcal{D}}^e$ and $P = P_+$ is the only Hermitian solution with the property that $\sigma(A_P) \subset \bar{\mathcal{D}}$.

THEOREM 3.2. *Assume (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Assume that* (2.9) *has at least one Hermitian solution. Let P_- and P_+ be the solutions determined by* (3.5). *Then $\sigma(A_+) \subset \bar{\mathcal{D}}$ and $\sigma(A_-) \subset \bar{\mathcal{D}}^e$. Furthermore, for any Hermitian solution of* (2.9), *we have*

$$P_- \leq P \leq P_+.$$

In addition, if P is a Hermitian solution with the property that $\sigma(A_P) \subset \bar{\mathcal{D}}^e$, then $P = P_-$. If P is a Hermitian solution with the property that $\sigma(A_P) \subset \bar{\mathcal{D}}$, then $P = P_+$.

Proof. Let P be a Hermitian solution. Then $X = P - P_-$ solves (3.6) and $S_- + B^*XB = R + B^*PB > 0$. We shall prove that $X \geq 0$. First we shall prove that $\mathcal{X}_0(A_-) \subseteq \ker X$. In order to prove this, choose a basis of $\mathcal{X}_0(A_-)$ consisting of eigenvectors and generalized eigenvectors. Such a basis consists of chains of vectors $x_1, \ldots, x_k$ with the property that

$$\begin{aligned} A_-x_1 &= \lambda x_1, \\ A_-x_2 &= \lambda x_2 + x_1, \\ &\vdots \\ A_-x_k &= \lambda x_k + x_{k-1}, \end{aligned}$$

where $|\lambda| = 1$. We will show by induction that $Xx_1 = \cdots = Xx_k = 0$. Assume $A_-x_1 = \lambda x_1$. Using (3.6) we obtain

$$x_1^*Xx_1 = |\lambda|^2 x_1^*Xx_1 - |\lambda|^2 x_1^*XB(S_- + B^*XB)^{-1}B^*Xx_1.$$

This yields $x_1^*XB(S_- + B^*XB)^{-1}B^*Xx_1 = 0$, from which we obtain $x_1^*XB = 0$. Again using (3.6), this implies $x_1^*X = \bar{\lambda}x_1^*XA_-$, from which

$$x_1^*X(A_- - \bar{\lambda}^{-1}I \quad B) = 0.$$

By controllability of (A, B), the latter implies that $Xx_1 = 0$. Now, assume that $A_- x_r = x_{r-1} + \lambda x_r$ and $Xx_{r-1} = 0$. Using (3.6) we find

$$x_r^* X x_r = (x_{r-1}^* + \bar{\lambda} x_r^*) X (x_{r-1} + \lambda x_r) \\ - (x_{r-1}^* + \bar{\lambda} x_r^*) X B (S_- + B^* X B)^{-1} B^* X (x_{r-1} + \lambda x_r).$$

The latter implies that $(x_{r-1}^* + \bar{\lambda} x_r^*) X B = 0$ and hence $x_r^* X B = 0$. Also, $x_r^* X = \bar{\lambda} x_r^* X A_-$. Again, by controllability of (A, B), this yields $Xx_r = 0$. This proves our claim that $\mathcal{X}_0(A_-) \subseteq \ker X$.

To proceed, let U be a unitary matrix such that

$$U^* A_- U = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

with $\sigma(A_{11}) \subset \mathcal{T}$ and $\sigma(A_{22}) \subset \mathcal{D}^e$. Let $Y = U^* X U$. Obviously, since $\mathcal{X}_0(A_-) \subseteq \ker X$, we have

$$Y = \begin{pmatrix} 0 & 0 \\ 0 & Y_{22} \end{pmatrix}.$$

Furthermore, it follows from (3.6) that $Y_{22} - A_{22}^* Y_{22} A_{22} \leq 0$. Since the eigenvalues of A_{22} lie *strictly* outside the unit disc, the latter can be shown to imply $Y_{22} \geq 0$ (see, e.g., [9, Thm. 13.2.3]). Thus we have proven $X \geq 0$.

A completely similar argument can be used to show that any Hermitian solution P satisfies $P \leq P_+$. Finally, note that in the above proof we only used the facts that $\sigma(A_-) \subset \bar{\mathcal{D}}^e$ and $\sigma(A_+) \subset \bar{\mathcal{D}}$. Hence, if P is a Hermitian solution with the property that $\sigma(A_P) \subset \bar{\mathcal{D}}^e$, then we must have $P \leq Q$ for any Hermitian solution Q, in particular, for $Q = P_-$. This shows that $P = P_-$. The statement on P_+ is proven similarly. □

Next we prove an analogue for the discrete-time case of a theorem first proved by Coppel [1] for the continuous-time case.

THEOREM 3.3. *Assume that (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some η on the unit circle. Let P_- and P_+ be the minimal and maximal Hermitian solutions of* (2.9)*, respectively. Put $\Delta := P_+ - P_-$ and*

$$A_- := A - B(R + B^* P_- B)^{-1}(C + B^* P_- A).$$

Then, for every A_--invariant subspace $\mathcal{V}$ of $\mathcal{X}_+(A_-)$, we have

$$\mathcal{C}^n = \mathcal{V} \oplus (\Delta \mathcal{V})^\perp. \tag{3.12}$$

Let $\pi_{\mathcal{V}}$ denote the projection onto $\mathcal{V}$ along $(\Delta \mathcal{V})^\perp$. For every A_--invariant subspace $\mathcal{V} \subseteq \mathcal{X}_+(A_-)$ the matrix P defined by

$$P = P_- \pi_{\mathcal{V}} + P_+(I - \pi_{\mathcal{V}}) \tag{3.13}$$

is a Hermitian solution of (2.9)*. Conversely, for every Hermitian solution P of* (2.9) *there exists a unique A_--invariant subspace $\mathcal{V}$ of $\mathcal{X}_+(A_-)$ such that* (3.13) *holds. This subspace $\mathcal{V}$ is equal to $\mathcal{V} = \mathcal{X}_+(A_P)$, where A_P is defined by*

$$A_P = A - B(R + B^* P B)^{-1}(C + B^* P A).$$

In addition, $\mathcal{X}_0(A_P) = \ker \Delta$ *and* $\mathcal{X}_-(A_P) = (\Delta\mathcal{V})^\perp \cap \mathcal{X}_-(A_+)$.

It will become clear in the proof that this result is actually little more than a reformulation of the last part of Theorem 3.1.

Proof. First we show that it suffices to prove the theorem for equation (3.6). Note that $X_- = 0$ is the minimal solution of (3.6) and $X_+ = P_+ - P_-$ is the maximal solution. Moreover, by straightforward computation,

$$\begin{aligned}(A_-)_X &:= A_- - B(S_- + B^*XB)^{-1}B^*XA_- \\ &= A_{X+P_-}.\end{aligned} \tag{3.14}$$

In particular, this means that $(A_-)_{X_-} = A_-$ and $(A_-)_{X_+} = A_+$, i.e., the A_- and A_+ matrices remain the same. Because of Lemma 3.1, P is a solution of (2.9) if and only if $P = X + P_-$ for some solution of (3.6). Now assume the theorem is true for (3.6). Let $\mathcal{V}$ be an A_--invariant subspace of $\mathcal{X}_+(A_-)$. As the A_- matrix remains the same, we conclude that (3.12) holds. Furthermore, the matrix $X := X_+(I - \pi_\mathcal{V})$ is a Hermitian solution of (3.6). This implies that $P := P_- + X = P_-\pi_\mathcal{V} + P_+(I - \pi_\mathcal{V})$ is a Hermitian solution of (2.9). Conversely, let P be a solution of (2.9). Define $X := P - P_-$. There exists an A_--invariant subspace of $\mathcal{X}_+(A_-)$ such that $X = X_+(I - \pi_\mathcal{V})$. This, however, yields $P = P_-\pi_\mathcal{V} - P_+(I - \pi_\mathcal{V})$. Finally, since $(A_-)_X = A_P$ and $(A_-)_{X_+} = A_+$, we also find $\mathcal{V} = \mathcal{X}_+(A_P)$, $\ker \Delta = \mathcal{X}_0(A_P)$, and $\mathcal{X}_-(A_+) \cap (\Delta\mathcal{V})^\perp = \mathcal{X}_-(A_P)$.

It remains to prove the statements of the theorem for equation (3.6). The matrix T given by (3.1) looks particularly simple in the case of equation (3.6):

$$T = \begin{pmatrix} A_- & -BS_-^{-1}B^*A_-^{*-1} \\ 0 & A_-^{*-1} \end{pmatrix},$$

with $S_- > 0$. Note that A_- has all its eigenvalues on or outside the unit circle. Hence we have $\mathcal{X}_+(T) \subseteq \operatorname{im}\binom{I}{0}$; more precisely,

$$\mathcal{X}_+(T) = \mathcal{X}_+(A_-) \times \{0\}.$$

So, T-invariant subspaces $\mathcal{N}$ with $\sigma(T \mid \mathcal{N}) \subset \mathcal{D}^e$ are precisely the subspaces of the form

$$\mathcal{N} = \mathcal{V} \times \{0\},$$

where $\mathcal{V}$ is A_--invariant and $\mathcal{V} \subseteq \mathcal{X}_+(A_-)$.

Now, let $\mathcal{V}$ be an A_--invariant subspace of $\mathcal{X}_+(A_-)$. We will prove (3.12) and the fact that $X_+(I - \pi_\mathcal{V})$ is a Hermitian solution of (3.6). According to Theorem 3.1, there is a unique solution X of (3.6) such that

$$\operatorname{im}\begin{pmatrix} I \\ X \end{pmatrix} \cap (\mathcal{X}_+(A_-) \times \{0\}) = \mathcal{V} \times \{0\}.$$

From [3, §§2, 7] and Theorem 3.1 we have

$$\operatorname{im}\begin{pmatrix} I \\ X \end{pmatrix} = (\mathcal{V} \times \{0\}) \oplus \left(\operatorname{im}\begin{pmatrix} I \\ X \end{pmatrix} \cap \mathcal{X}_0(T)\right) \oplus \left(\mathcal{X}_-(T) \cap \left(\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\mathcal{N}\right)^\perp\right),$$

and moreover $\operatorname{im}\binom{I}{X} \cap \mathcal{X}_0(T)$ is the same subspace for any Hermitian solution X (here we also use the fact that the signs in the sign characteristic of

$$\left(T, \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\right)$$

are all the same; see [4, Thm. 1.2]). Then

$$\begin{aligned}\operatorname{im}\begin{pmatrix} I \\ X \end{pmatrix} \cap \mathcal{X}_0(T) &= \operatorname{im}\begin{pmatrix} I \\ X_+ \end{pmatrix} \cap \mathcal{X}_0(T) \\ &= \operatorname{im}\begin{pmatrix} I \\ X_- \end{pmatrix} \cap \mathcal{X}_0(T) \\ &= \operatorname{im}\begin{pmatrix} I \\ 0 \end{pmatrix} \cap \mathcal{X}_0(T) \\ &\subseteq \ker X_+ \times \{0\}.\end{aligned}$$

Conversely, for $x \in \ker X_+$, we have from (3.6)

$$0 = A^*{}_- X_+(A_- - B(S_- + B^* X_+ B)^{-1} B^* X_+ A_-)x = A^*{}_- X_+ A_+ x = 0.$$

So $X_+ A_+ x = 0$, i.e., $\ker X_+$ is A_+-invariant. Consider the equality

$$T\begin{pmatrix} I \\ X_+ \end{pmatrix} x = \begin{pmatrix} I \\ X_+ \end{pmatrix} A_+ x,$$

which holds by Theorem 3.1, in particular, the first description of the set of solutions. For $x \in \ker X_+$ this gives

$$T\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} A_+ x \\ 0 \end{pmatrix}.$$

However,

$$T\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} A_- x \\ 0 \end{pmatrix}.$$

It follows that for $x \in \ker X_+$ we have $A_+ x = A_- x$, and hence $\ker X_+$ is also A_--invariant. We claim that

$$\ker X_+ \subseteq \mathcal{X}_0(A_-). \tag{3.15}$$

Indeed, assume $x \in \ker X_+$, $x \neq 0$, and $A_- x = \lambda x$. Then $T\binom{x}{0} = \lambda\binom{x}{0}$. Since

$$\begin{pmatrix} x \\ 0 \end{pmatrix} \in \operatorname{im}\begin{pmatrix} I \\ X_+ \end{pmatrix},$$

by Lemma 3.3 we have $\lambda \in \bar{\mathcal{D}}$. On the other hand, since $\lambda \in \sigma(A_-)$, $\lambda \in \bar{\mathcal{D}}^e$. Thus $\lambda \in \mathcal{T}$, which proves the claim. It follows from (3.15) that

$$\operatorname{im}\begin{pmatrix} I \\ X \end{pmatrix} \cap \mathcal{X}_0(T) = \ker X_+ \times \{0\}.$$

Next,

$$\mathcal{X}_-(T) = \mathcal{X}_-(T) \cap \operatorname{im}\begin{pmatrix} I \\ X_+ \end{pmatrix},$$

so

$$\mathcal{X}_-(T) \cap \left(\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\mathcal{N}\right)^{\perp} \subseteq \operatorname{im}\begin{pmatrix} I \\ X_+ \end{pmatrix} \cap (\mathcal{C}^n \times \mathcal{V}^{\perp}).$$

Now

$$\begin{pmatrix} x \\ X_+ x \end{pmatrix} \in \mathcal{C}^n \times \mathcal{V}^\perp$$

implies $X_+ x \in \mathcal{V}^\perp$, i.e., $x \in (X_+\mathcal{V})^\perp$. So

$$\mathcal{X}_-(T) \cap \left(\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \mathcal{N} \right)^\perp \subseteq \left\{ \begin{pmatrix} x \\ X_+ x \end{pmatrix} \mid x \in (X_+\mathcal{V})^\perp \right\}.$$

Since obviously $\ker X_+ \subseteq (X_+\mathcal{V})^\perp$ also, we find

$$\text{im} \begin{pmatrix} I \\ X \end{pmatrix} \subseteq (\mathcal{V} \times \{0\}) + \left\{ \begin{pmatrix} x \\ X_+ x \end{pmatrix} \mid x \in (X_+\mathcal{V})^\perp \right\}. \tag{3.16}$$

Using the previous inclusion, it is easy to see that $\mathcal{C}^n = \mathcal{V} + (X_+\mathcal{V})^\perp$. We claim that the latter is, in fact, a direct sum. Indeed, $x \in \mathcal{V} \cap (X_+\mathcal{V})^\perp$ implies $(x, X_+ x) = 0$, from which $X_+ x = 0$. Thus $\mathcal{V} \cap (X_+\mathcal{V})^\perp \subseteq \ker X_+ \cap \mathcal{V} = 0$ (recall that $\mathcal{V} \subseteq \mathcal{X}_+(A_-)$ while $\ker X_+ \subseteq \mathcal{X}_0(A_-)$). Now, if $\pi_\mathcal{V}$ is the projection onto $\mathcal{V}$ along $(X_+\mathcal{V})^\perp$, it can be seen that (3.16) implies

$$X = X_+(I - \pi_\mathcal{V}).$$

Next, conversely, let X be a Hermitian solution of (3.10). Define

$$\mathcal{N} := \text{im} \begin{pmatrix} I \\ X \end{pmatrix} \cap \mathcal{X}_+(T).$$

Then $\mathcal{N}$ is a T-invariant subspace and $\sigma(T \mid \mathcal{N}) \subset \mathcal{D}^e$. Thus $\mathcal{N}$ has the form $\mathcal{N} = \mathcal{V} \times \{0\}$ for some A_--invariant subspace $\mathcal{V}$ of $\mathcal{X}_+(A_-)$. By repeating the argument in the first part of this proof, it is then shown that we must have $X = X_+(I - \pi_\mathcal{V})$.

Finally, we will show that if $\mathcal{V} \subset \mathcal{X}_+(A_-)$ is A_--invariant and $X = X_+(I - \pi_\mathcal{V})$, then $\mathcal{V} = \mathcal{X}_+((A_-)_X)$, $\ker X_+ = \mathcal{X}_0((A_-)_X)$, and $(\Delta\mathcal{V})^\perp \cap \mathcal{X}_-(A_+) = \mathcal{X}_-((A_-)_X)$.

As in the proof of Theorem 3.2, we show that $\mathcal{X}_0(A_-) \subseteq \ker X_+$. Combined with (3.15) this yields $\ker X_+ = \mathcal{X}_0(A_-)$. Since $0 \leq X \leq X_+$, $\ker X_+ \subseteq \ker X$. Hence $(A_-)_X \mid \ker X_+ = A_- \mid \ker X_+$. This yields

$$\ker X_+ = \mathcal{X}_0((A_-)_X),$$

as desired. We will now show that

$$\mathcal{V} \subseteq \mathcal{X}_+((A_-)_X). \tag{3.17}$$

Indeed, note that $\mathcal{V} \subseteq \ker X$. Hence $(A_-)_X \mid \mathcal{V} = A_- \mid \mathcal{V}$, which yields (3.17). Next, we show that

$$(X_+\mathcal{V})^\perp \subseteq \mathcal{X}_0((A_-)_X) \oplus \mathcal{X}_-((A_-)_X). \tag{3.18}$$

In order to prove this, first note that $A_-\mathcal{V} = \mathcal{V}$, since $\mathcal{V}$ is A_--invariant and since A_- is invertible. Now, the Riccati equation (3.6) with $X = X_+$ can be written as $X_+ = A_-^* X_+ A_+$. We claim that $(X_+\mathcal{V})^\perp$ is A_+-invariant. Let $w \in (X_+\mathcal{V})^\perp$. Then for all $v \in \mathcal{V}$,

$$0 = v^* X_+ w = v^* A_-^* X_+ A_+ w.$$

Hence $A_+w \in (X_+A_-\mathcal{V})^\perp = (X_+\mathcal{V})^\perp$. Next, by straightforward calculation, we show that

$$(A_-)_X = A_+ - B(R + B^*(X + P_-)B)^{-1}B^*(X - X_+)A_+. \tag{3.19}$$

Since $X \mid (X_+\mathcal{V})^\perp = X_+ \mid (X_+\mathcal{V})^\perp$ (3.19) yields $(A_-)_X \mid (X_+\mathcal{V})^\perp = A_+ \mid (X_+\mathcal{V})^\perp$. Since $\sigma(A_+) \subset \bar{\mathcal{D}}$, this implies $\sigma((A_-)_X \mid (X_+\mathcal{V})^\perp) \subset \bar{\mathcal{D}}$, which yields (3.18).

By combining (3.17), (3.18), and the fact that $\mathcal{V} \oplus (X_+\mathcal{V})^\perp = \mathcal{C}^n$, we find that the inclusions (3.17) and (3.18) are, in fact, equalities. Finally, since $(A_-)_X \mid (X_+\mathcal{V})^\perp = A_+ \mid (X_+\mathcal{V})^\perp$, we find $(X_+\mathcal{V})^\perp \cap \mathcal{X}_-(A_+) = \mathcal{X}_-((A_-)_X)$. □

4. A proof of Theorem 2.1. The proof is split up into several lemmas, which are all discrete-time counterparts of results in [6]. In this section, let $\mathcal{L}$ be an arbitratry but fixed subspace of $\mathcal{C}^n$. We will first study the finiteness of the optimal cost $V_{\mathcal{L}}(x_0)$. Note that our assumption that (A, B) is controllable is sufficient to guarantee that $V_{\mathcal{L}}(x_0) < +\infty$ for all x_0. In the sequel we shall establish a sufficient condition to guarantee $V_{\mathcal{L}}(x_0) > -\infty$ for all x_0. From §2 recall the definition of negative semidefiniteness on $\mathcal{L}$ of a given Hermitian matrix P. It turns out that if the smallest solution P_- of the Riccati equation (2.9) is negative semidefinite on $\mathcal{L}$, then the optimal cost is finite.

LEMMA 4.1. *Assume that (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Furthermore, assume that* (2.9) *has at least one Hermitian solution. Then we have: if P_- is negative semidefinite on $\mathcal{L}$, then $V_{\mathcal{L}}(x_0) \in \mathcal{R}$ for all $x_0 \in \mathcal{C}^n$.*

Our proof of Lemma 4.1 uses the following two lemmas.

LEMMA 4.2. *Let $\mathcal{L}$ be a subspace of $\mathcal{C}^n$ and let H be a matrix such that $\mathcal{L} = \ker H$. Let $P \in \mathcal{C}^{n\times n}$ be Hermitian. Then P is negative semidefinite on $\mathcal{L}$ if and only if there exists $\lambda \in \mathcal{R}$ such that $P - \lambda H^*H$ is negative semidefinite.*

For a proof of the above lemma, we refer to [6].

LEMMA 4.3. *For any $x_0 \in \mathcal{C}^n$, any sequence u and any Hermitian solution P of* (2.9) *we have*

$$J_T(x_0, u) = x_0^*Px_0 - x_{T+1}^*Px_{T+1} + \sum_{k=0}^{T} \|u_k + (R + B^*PB)^{-1}(C + B^*PA)x_k\|^2_{R+B^*PB}.$$

Here, $\|v\|_S^2 := v^*Sv$.

Proving this lemma is just a matter of standard computation.

Proof of Lemma 4.1. Let $x_0 \in \mathcal{C}^n$. Since (A, B) is controllable, there is an input sequence $u \in U_{\mathcal{L}}(x_0)$ such that $J(x_0, u) < +\infty$ (in fact, one can steer from x_0 to the origin in finite time). Thus $V_{\mathcal{L}}(x_0) \in \mathcal{R} \cup \{-\infty\}$. Now let $u \in U_{\mathcal{L}}(x_0)$ be arbitrary. Let H be such that $\mathcal{L} = \ker H$ and let $\lambda \in \mathcal{R}$ be such that $P_- - \lambda H^*H$ is negative semidefinite. According to Lemma 4.3, for all T we have

$$J_T(x_0, u) = x_0^*P_-x_0 - x_{T+1}^*(P_- - \lambda H^*H)x_{T+1} - \lambda\|Hx_{T+1}\|^2 + \sum_{k=0}^{T} \|u_k + (R + B^*PB)^{-1}(C + B^*PA)x_k\|^2_{R+B^*PB}.$$

Thus, for all $T \geq 0$,

$$J_T(x_0, u) \geq x_0^*P_-x_0 - \lambda\|Hx_{T+1}\|^2.$$

Since x_T converges to $\mathcal{L}$ as $T \to \infty$, we have $Hx_{T+1} \to 0$. It follows that

$$J(x_0, u) = \lim_{T\to\infty} J_T(x_0, u) \geq x_0^* P_- x_0.$$

Since the latter holds for all $u \in U_{\mathcal{L}}(x_0)$ this proves our claim. $\square$

The next few lemmas give some general properties of linear systems.

LEMMA 4.4. *Consider the system*

$$x_{k+1} = Ax_k + v_k, \qquad y_k = Cx_k,$$

and suppose that (C, A) *is observable. Then if* $\{v_k\}_{k=0}^{\infty} \in \ell_2$ *and* $\{y_k\}_{k=0}^{\infty} \in \ell_\infty$, *necessarily* $\{x_k\}_{k=0}^{\infty} \in \ell_\infty$.

Proof. Since (C, A) is observable, there exists a matrix L such that $\sigma(A + LC) \subset \mathcal{D}$. Obviously, $\{x_k\}_{k=0}^{\infty}$ satisfies the difference equation

$$x_{k+1} = (A + LC)x_k - Ly_k + v_k.$$

Using some straightforward estimates we see that $v \in \ell_2$ and $y \in \ell_\infty$ imply $x \in \ell_\infty$. $\square$

In the following lemma, if $\mathcal{C}_g$ is a subset of $\mathcal{C}$, then $\mathcal{X}_g(A)$ will denote the spectral subspace of A associated with its eigenvalues in $\mathcal{C}_g$, i.e., the largest A-invariant subspace $\mathcal{V}$ with the property that $\sigma(A \mid \mathcal{V}) \subset \mathcal{C}_g$. Using the previous lemma, we can now prove the following.

LEMMA 4.5. *Consider the system*

$$x_{k+1} = Ax_k + v_k, \qquad y_k = Cx_k.$$

Assume that (C, A) *is detectable (relative to* $\mathcal{C}_g$*). Let the state space* $\mathcal{C}^n$ *be decomposed into* $\mathcal{C}^n = \mathcal{X}_1 \oplus \mathcal{X}_2$, *where* $\mathcal{X}_1$ *is* A*-invariant. In this decomposition, let*

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Assume that $\sigma(A \mid \mathcal{X}_1) \subset \mathcal{C}_g$ *and* $\sigma(A \mid \mathcal{C}^n/\mathcal{X}_1) \subset \mathcal{C}/\mathcal{C}_g$. *Then for every initial condition* x_0 *we have: if* $\{v_k\}_{k=0}^{\infty} \in \ell_2$ *and* $\{y_k\}_{k=0}^{\infty} \in \ell_\infty$, *then* $\{x_{2,k}\}_{k=0}^{\infty} \in \ell_\infty$.

Proof. We claim that $\mathcal{X}_1 = \mathcal{X}_g(A)$. Indeed, by assumption, we have $\mathcal{X}_1 \subseteq \mathcal{X}_g(A)$. Denote $\sigma_0 := \sigma(A \mid \mathcal{X}_g(A)/\mathcal{X}_1)$. Then we have $\sigma_0 \subset \mathcal{C}_g$. Also, $\sigma_0 \subset \sigma(A \mid \mathcal{C}^n/\mathcal{X}_1) \subset \mathcal{C}/\mathcal{C}_g$. This implies that $\sigma_0 = \emptyset$ or, equivalently, $\mathcal{X}_1 = \mathcal{X}_g(A)$. By the fact that (C, A) is detectable with respect to $\mathcal{C}_g$, we may now conclude that $\langle \ker C \mid A \rangle \subseteq \mathcal{X}_1$. Decompose $\mathcal{X}_1 = \mathcal{X}_{11} \oplus \mathcal{X}_{12}$, with $\mathcal{X}_{11} := \langle \ker C \mid A \rangle$ and $\mathcal{X}_{12}$ arbitrary. Accordingly, partition

$$x_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}.$$

We then have $\mathcal{C}^n = \mathcal{X}_{11} \oplus \mathcal{X}_{12} \oplus \mathcal{X}_2$ with $x = (x_{11}^T, x_{12}^T, x_2^T)^T$. In this decomposition let

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix}, \quad C = \begin{pmatrix} 0 & C_2 & C_3 \end{pmatrix}, \quad v = \begin{pmatrix} v_{11} \\ v_{12} \\ v_2 \end{pmatrix}.$$

Obviously, the system

$$\left(\begin{pmatrix} C_2 & C_3 \end{pmatrix}, \begin{pmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{pmatrix} \right)$$

is observable. Moreover,

$$\begin{pmatrix} x_{12,k+1} \\ x_{2,k+1} \end{pmatrix} = \begin{pmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{pmatrix} \begin{pmatrix} x_{12,k} \\ x_{2,k} \end{pmatrix} + \begin{pmatrix} v_{12,k} \\ v_{2,k} \end{pmatrix},$$

$$y_k = \begin{pmatrix} C_2 & C_3 \end{pmatrix} \begin{pmatrix} v_{12,k} \\ v_{2,k} \end{pmatrix}.$$

It now follows from the previous lemma that

$$\begin{pmatrix} x_{12} \\ x_2 \end{pmatrix} \in \ell_\infty,$$

which implies that $x_2 \in \ell_\infty$. □

The next lemma tells us that a semistable controllable system has the property that all initial states can be steered to the origin with arbitrary small controls (in ℓ_2-sense).

LEMMA 4.6. *Consider the controllable system*

$$x_{k+1} = Ax_k + Bu_k, \quad x_0 \text{ given.}$$

Assume that $\sigma(A) \subset \bar{\mathcal{D}}$. Then for all $\epsilon > 0$ there exists a $u \in \ell_2$ such that $\|u\|_2 < \epsilon$ and $x(x_0, u)_k \to 0$ as $k \to \infty$.

Proof. We will first show that it suffices to prove the statement of the lemma with $\bar{\mathcal{D}}$ replaced by $\mathcal{T}$. Indeed, we can always choose a basis in $\mathcal{C}^n$ such that A and B have the form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

with $\sigma(A_1) \subset \mathcal{D}$ and $\sigma(A_2) \subset \mathcal{T}$. Consider the subsystem $x_{1,k+1} = A_1 x_{1,k} + Bu_k$, with $x_{1,0}$ given. Obviously, for any input sequence $u \in \ell_2$, we automatically have that $x_1 \in \ell_2$ and hence $x_{1,k} \to 0$ as $k \to \infty$.

Assume, therefore, that $\sigma(A) \subset \mathcal{T}$. For any $\delta > 0$, any initial state x_0 and any input sequence u, define the quadratic cost

$$J_\delta(x_0, u) = \sum_{k=0}^{\infty} \|u_k\|^2 + \delta^2 \|x(x_0, u)_k\|^2 \tag{4.1}$$

and consider the optimization problem

$$\inf_u J_\delta(x_0, u). \tag{4.2}$$

Note that (4.2) can be considered as a "standard" discrete-time linear quadratic problem of minimizing $\sum_{k=0}^{\infty} u_k^* R u_k + x_k^* D^* D x_k$, with $R = I$ and $D = \delta I$ (see, for example, [7]). Since (A, B) is controllable and $(\delta I, A)$ is observable, the infimum (4.2) is equal to $x_0^* P(\delta) x_0$, with $P(\delta)$ the unique positive semidefinite solution of the Riccati equation

$$P = A^* P A + \delta^2 I - A^* P B (I + B^* P B)^{-1} B^* P A.$$

(In fact, $P(\delta) > 0$.) Furthermore, for any x_0 there exists a unique optimal input u^+ which is given by the state feedback control law

$$u^+ = -(I + B^*P(\delta)B)^{-1}B^*P(\delta)Ax^+,$$

and the corresponding optimal closed loop matrix

$$A - B(I + B^*P(\delta)B)^{-1}B^*P(\delta)A$$

is stable, i.e., has all its eigenvalues in $\mathcal{D}$. Now, we shall analyze what happens if $\delta \downarrow 0$. We claim that $P(\delta) \downarrow 0$. Indeed, $P(\delta) \geq 0$ and $P(\delta)$ is monotonically decreasing as $\delta \downarrow 0$. Hence there exists a Hermitian matrix $\bar{P} \geq 0$ such that $P(\delta) \downarrow \bar{P}$. Clearly, $\bar{P}$ satisfies the Riccati equation

$$P = A^*PA - A^*PB(I + B^*PB)^{-1}B^*PA. \tag{4.3}$$

Now, by applying Theorem 3.2 we find that the latter Riccati equation has a largest Hermitian solution, say, P_+. We contend that $P_+ = 0$. Indeed, $P = 0$ is a solution of (4.3) and it has the property that $\sigma(A_P) \subset \bar{\mathcal{D}}$ (since $A_P = A$ and $\sigma(A) \subset \mathcal{T}$). Thus we must have $\bar{P} \leq 0$. Our conclusion is that $\bar{P} = 0$.

In order to complete the proof, let $\epsilon > 0$. Choose $\delta > 0$ such that $x_0^*P(\delta)x_0 < \epsilon$. Choose the input sequence given by $u^+ = -(I + B^*P(\delta)B)B^*P(\delta)Ax^+$. Then we have

$$\|u^+\|^2 \leq J_\delta(x_0, u^+) = x_0^*P(\delta)x_0 < \epsilon.$$

Since the corresponding closed loop matrix is stable, the corresponding state trajectory converges to zero as $k \to \infty$. $\square$

We proceed by decomposing $\mathcal{C}^n$ as follows. Let

$$\mathcal{X}_1 := \langle \mathcal{L} \cap \ker P_- \mid A_- \rangle \cap \mathcal{X}_+(A_-) = \mathcal{V}(\mathcal{L}),$$

and let $P_\mathcal{L}$ be the solution corresponding to $\mathcal{V}(\mathcal{L})$ according to Theorem 3.3. Put

$$A_\mathcal{L} := A - B(R + B^*P_\mathcal{L}B)^{-1}(C + B^*P_\mathcal{L}A).$$

According to Theorem 3.3, we have $\mathcal{X}_+(A_\mathcal{L}) = \mathcal{X}_1$. In addition, we define

$$\mathcal{X}_2 := \mathcal{X}_0(A_\mathcal{L}) = \mathcal{X}_0(A_-) = \ker \Delta,$$
$$\mathcal{X}_3 := \mathcal{X}_-(A_\mathcal{L}) = (\Delta\mathcal{V}(\mathcal{L}))^\perp \cap \mathcal{X}_-(A_+).$$

With respect to the decomposition $\mathcal{C}^n = \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_3$, we have

$$A_- = \begin{pmatrix} A_{11} & 0 & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix}, \qquad A_\mathcal{L} = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A'_{33} \end{pmatrix}.$$

Here, we have used that $A_- \mid \mathcal{X}_1 = A_\mathcal{L} \mid \mathcal{X}_1$ and that $A_- \mid \mathcal{X}_2 = A_\mathcal{L} \mid \mathcal{X}_2$. This follows most easily by combining (3.14) and the facts that $P_\mathcal{L} \mid \mathcal{V}(\mathcal{L}) = P_- \mid \mathcal{V}(\mathcal{L})$ and $P_\mathcal{L} \mid (\Delta\mathcal{V}(\mathcal{L}))^\perp = P_+ \mid (\Delta\mathcal{V}(\mathcal{L}))^\perp$. We also have

$$P_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & P_{22}^- & P_{23}^- \\ 0 & P_{23}^{-*} & P_{33}^- \end{pmatrix}, \qquad \Delta = \begin{pmatrix} \Delta_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta_{33} \end{pmatrix}$$

and

$$P_+ = P_- + \Delta = \begin{pmatrix} \Delta_{11} & 0 & 0 \\ 0 & P_{22}^- & P_{23}^- \\ 0 & P_{23}^{-*} & P_{33}^+ \end{pmatrix}, \qquad P_{\mathcal{L}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & P_{22}^- & P_{23}^- \\ 0 & P_{23}^{-*} & P_{33}^+ \end{pmatrix},$$

where $P_{33}^+ := P_{33}^- + \Delta_{33}$.

The next lemma states that $P_{\mathcal{L}}$ yields a lower bound for the linear quadratic problem (2.8) under consideration.

LEMMA 4.7. *Suppose (A, B) is controllable, R is nonsingular, $A - BR^{-1}C$ is nonsingular, and $\Psi(\eta) > 0$ for some $\eta \in \mathcal{T}$. Assume further that (2.9) has at least one Hermitian solution and assume that P_- is negative semidefinite on $\mathcal{L}$. Then for all x_0 and for all $u \in U_{\mathcal{L}}(x_0)$ we have*

$$J(x_0, u) \geq x_0^* P_{\mathcal{L}} x_0 + \sum_{k=0}^{\infty} \|u_k + (R + B^* P_{\mathcal{L}} B)^{-1}(C + B^* P_{\mathcal{L}} A) x_k\|^2_{R+B^* P_{\mathcal{L}} B}.$$

Proof. Let H be a matrix such that $L = \ker H$. Let $\lambda \in \mathcal{R}$ be such that $P_- - \lambda H^* H \leq 0$ (see Lemma 4.2). Take an arbitrary $u \in U_{\mathcal{L}}(x_0)$. It follows from Lemma 4.1 that $J(x_0, u) \in \mathcal{R} \cup \{+\infty\}$. If it is equal to $+\infty$, then the inequality trivially holds. Assume therefore that $J(x_0, u)$ is finite. Put

$$v_k = u_k + (R + B^* P_- B)^{-1}(C + B^* P_- A) x_k.$$

From Lemma 4.3 we have

(4.4)
$$\begin{aligned} \sum_{k=0}^{T} \|v_k\|^2_{R+B^* P_- B} &= J_T(x_0, u) - x_0^* P_- x_0 + x_{T+1}^*(P_- - \lambda H^* H) x_{T+1} + \lambda \|H x_{T+1}\|^2 \\ &\leq J_T(x_0, u) - x_0^* P_- x_0 + \lambda \|H x_{T+1}\|^2. \end{aligned}$$

Since $J_T(x_0, u) \to J(x_0, u)$ and $Hx_T \to 0$, we find that $\{v_k\} \in \ell_2$. Again, using (4.4), this implies that

$$\lim_{T \to \infty} x_T^*(P_- - \lambda H^* H) x_T$$

exists and is finite. Thus $\lim_{T\to\infty} x_T^* P_- x_T$ exists and is finite. Also, since $P_- - \lambda H^* H$ is semidefinite, $(P_- - \lambda H^* H) x_k$ and hence $P_- x_k$ are bounded functions of k. Denote

$$y_k = \begin{pmatrix} P_- \\ H \end{pmatrix} x_k.$$

Then $\{y_k\} \in \ell_\infty$. Since $x_{k+1} = Ax_k + Bu_k$, we have that $\{x_k\}$, $\{y_k\}$, and $\{v_k\}$ are related by

$$x_{k+1} = A_- x_k + B v_k, \qquad y_k = \begin{pmatrix} P_- \\ H \end{pmatrix} x_k.$$

Now decompose $\mathcal{C}^n$ as above: $\mathcal{C}^n = \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_3$. Since

$$\mathcal{X}_1 \subseteq \ker \begin{pmatrix} P_- \\ H \end{pmatrix},$$

we have

$$\begin{pmatrix} P_- \\ H \end{pmatrix} = \begin{pmatrix} 0 & D_2 & D_3 \end{pmatrix} \tag{4.5}$$

for given matrices D_2 and D_3. With respect to the given decomposition, let

$$B = \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}.$$

Since $\mathcal{X}_1$ is the undetectable subspace (relative to $\bar{\mathcal{D}}$) of the system

$$\left(\begin{pmatrix} P_- \\ H \end{pmatrix}, A_- \right),$$

it is easily verified that the pair

$$\left(\begin{pmatrix} D_2 & D_3 \end{pmatrix}, \begin{pmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{pmatrix} \right)$$

is detectable (relative to $\bar{\mathcal{D}}$). Since $\sigma(A_-) \subset \bar{\mathcal{D}}^e$ and $\mathcal{X}_2 = \mathcal{X}_0(A_-)$, we have $\sigma(A_{22}) \subset \mathcal{T}$ and

$$\sigma\left(\begin{pmatrix} A_{11} & A_{13} \\ 0 & A_{33} \end{pmatrix} \right) \subset \mathcal{D}^e.$$

Hence $\sigma(A_{33}) \subset \mathcal{D}^e$. Also we have

$$\begin{pmatrix} x_{2,k+1} \\ x_{3,k+1} \end{pmatrix} = \begin{pmatrix} A_{22} & A_{23} \\ 0 & A_{33} \end{pmatrix} \begin{pmatrix} x_{2,k} \\ x_{3,k} \end{pmatrix} + \begin{pmatrix} B_2 \\ B_3 \end{pmatrix} v_k,$$
$$y_k = \begin{pmatrix} D_2 & D_3 \end{pmatrix} \begin{pmatrix} x_{2,k} \\ x_{3,k} \end{pmatrix}.$$

Since $\{v_k\} \in \ell_2$ and $\{y_k\} \in \ell_\infty$, by Lemma 4.5 (applied with $\mathcal{C}_g = \bar{\mathcal{D}}$), we have $\{x_{3,k}\} \in \ell_\infty$. Now consider Lemma 4.3. We have

$$J_T(x_0, u) = x_0^* P_{\mathcal{L}} x_0 - x_{T+1}^* P_{\mathcal{L}} x_{T+1} + \sum_{k=0}^{T} \|w_k\|^2_{R+B^*P_{\mathcal{L}}B},$$

where

$$w_k = u_k + (R + B^* P_{\mathcal{L}} B)^{-1}(C + B^* P_{\mathcal{L}} A) x_k.$$

Then

$$\begin{aligned} J_T(x_0, u) &= x_0^* P_{\mathcal{L}} x_0 + \sum_{k=0}^{T} \|w_k\|^2_{R+B^*P_{\mathcal{L}}B} \\ &\quad - \begin{pmatrix} x_{2,T+1} \\ x_{3,T+1} \end{pmatrix}^* \begin{pmatrix} P_{22}^- & P_{23}^- \\ P_{23}^{-*} & P_{33}^+ \end{pmatrix} \begin{pmatrix} x_{2,T+1} \\ x_{3,T+1} \end{pmatrix} \\ &= x_0^* P_{\mathcal{L}} x_0 + \sum_{k=0}^{T} \|w_k\|^2_{R+B^*P_{\mathcal{L}}B} \\ &\quad - x_{3,T+1}^* \Delta_{33} x_{3,T+1} - x_{T+1}^* P_- x_{T+1}. \end{aligned} \tag{4.6}$$

Now $x^*_{T+1}P_-x_{T+1}$, $J_T(x_0,u)$ are bounded as $T\to\infty$, and likewise $x_{3,T+1}$ is bounded as $T\to\infty$. It follows that

$$\sum_{k=0}^{\infty}\|w_k\|^2_{R+B^*P_{\mathcal{L}}B}<\infty,$$

and as $R+B^*P_{\mathcal{L}}B\geq R+B^*P_-B>0$ we have $\{w_k\}\in\ell_2$. Now considering

$$x_{k+1}=Ax_k+Bu_k=A_{\mathcal{L}}x_k+Bw_k,$$

we see that $x_{3,k+1}=A'_{33}x_{3,k}+B_3w_k$. As $\sigma(A'_{33})\subset\mathcal{D}$ (because of the fact that $\mathcal{X}_3=\mathcal{X}_-(A_{\mathcal{L}})$) we obtain $\{x_{3,k}\}\in\ell_2$. Hence $\lim_{k\to\infty}x_{3,k}=0$. Now from (4.6) we have

$$\begin{aligned}J_T(x_0,u)&=x_0^*P_{\mathcal{L}}x_0-x^*_{T+1}(P_--\lambda H^*H)x_{T+1}\\&\quad-\lambda\|Hx_{T+1}\|^2+\sum_{k=0}^{T}\|w_k\|^2_{R+B^*P_{\mathcal{L}}B}-x^*_{3,T+1}\Delta_{33}x_{3,T+1}\\&\geq x_0^*P_{\mathcal{L}}x_0-\lambda\|Hx_{T+1}\|^2+\sum_{k=0}^{T}\|w_k\|^2_{R+B^*P_{\mathcal{L}}B}\\&\quad-x^*_{3,T+1}\Delta_{33}x_{3,T+1}.\end{aligned}$$

The desired result then follows by taking the limit as $T\to\infty$ in the above inequality. □

Our next lemma states that $V_{\mathcal{L}}(x_0)=x_0^*P_{\mathcal{L}}x_0$, taking into account the previous lemma.

LEMMA 4.8. *Suppose (A,B) is controllable, R is nonsingular, $A-BR^{-1}C$ is nonsingular, and $\Psi(\eta)>0$ for some $\eta\in\mathcal{T}$. Assume further that (2.9) has at least one Hermitian solution and assume that P_- is negative semidefinite on $\mathcal{L}$. Then for all x_0 and for all $\epsilon>0$, there is a $u\in U_{\mathcal{L}}(x_0)$ such that $J(x_0,u)\leq x_0^*P_{\mathcal{L}}x_0+\epsilon$.*

Proof. Put $w_k=u_k+(R+B^*P_{\mathcal{L}}B)^{-1}(C+B^*P_{\mathcal{L}}A)x_k$. From (4.6) we have for all $u\in U_{\mathcal{L}}(x_0)$

$$\begin{aligned}J_T(x_0,u)&=x_0^*P_{\mathcal{L}}x_0+\sum_{k=0}^{T}\|w_k\|^2_{R+B^*P_{\mathcal{L}}B}\\&\quad-\begin{pmatrix}x_{2,T+1}\\x_{3,T+1}\end{pmatrix}^*\begin{pmatrix}P_{22}^-&P_{23}^-\\P_{23}^{-*}&P_{33}^+\end{pmatrix}\begin{pmatrix}x_{2,T+1}\\x_{3,T+1}\end{pmatrix}.\end{aligned}$$

Moreover, $x_{k+1}=A_{\mathcal{L}}x_k+Bw_k$, so

$$\begin{pmatrix}x_{2,k+1}\\x_{3,k+1}\end{pmatrix}=\begin{pmatrix}A_{22}&0\\0&A'_{33}\end{pmatrix}\begin{pmatrix}x_{2,k}\\x_{3,k}\end{pmatrix}+\begin{pmatrix}B_2\\B_3\end{pmatrix}w_k.$$

Now $\sigma(A_{22})\subset\mathcal{T}$, $\sigma(A'_{33})\subset\mathcal{D}$. By Lemma 4.6 there is $w\in\ell_2$ such that

$$\sum_{k=0}^{\infty}\|w_k\|^2_{R+B^*P_{\mathcal{L}}B}<\epsilon\quad\text{and}\quad\begin{pmatrix}x_2\\x_3\end{pmatrix}(x_0,w)_k\to0\quad\text{as }k\to\infty.$$

Introduce

$$u_k=w_k-(R+B^*P_{\mathcal{L}}B)^{-1}(C+B^*P_{\mathcal{L}}A)x_k.$$

Then

$$\begin{aligned} J(x_0,u) &= \lim_{T\to\infty} J_T(x_0,u) \\ &= x_0^* P_{\mathcal{L}} x_0 + \sum_{k=0}^{\infty} \|w_k\|^2_{R+B^*P_{\mathcal{L}}B} \\ &\le \epsilon + x_0^* P_{\mathcal{L}} x_0. \qquad \square \end{aligned}$$

We now turn to the proof of Theorem 2.1.

Proof of Theorem 2.1. Part (i) is proved in Lemmas 4.1, 4.7, and 4.8. It remains to prove (ii).

(ii) First assume that for all x_0 there is an optimal control, i.e., $u^+ \in U_{\mathcal{L}}(x_0)$, for which $V_{\mathcal{L}}(x_0) = J(x_0,u^+)$. Choose x_0 and let u^+ be the corresponding optimal control. Put $x_k^+ = x(x_0,u^+)_k$. By Lemma 4.7

$$x_0^* P_{\mathcal{L}} x_0 = J(x_0,u^+) \ge x_0^* P_{\mathcal{L}} x_0 + \sum_{k=0}^{\infty} \|w_k\|^2_{R+B^*P_{\mathcal{L}}B},$$

where $w_k = u_k^+ + (R+B^*P_{\mathcal{L}}B)^{-1}(C+B^*P_{\mathcal{L}}A)x_k^+$. Hence $w_k = 0$, i.e.,

$$u_k^+ = -(R+B^*P_{\mathcal{L}}B)^{-1}(C+B^*P_{\mathcal{L}}A)x_k^+. \tag{4.7}$$

Since $x_{k+1} = A_{\mathcal{L}}x_k + Bw_k$, this yields $x_{k+1}^+ = A_{\mathcal{L}}x_k^+$; in particular,

$$\begin{pmatrix} x_{2,k+1}^+ \\ x_{3,k+1}^+ \end{pmatrix} = \begin{pmatrix} A_{22} & 0 \\ 0 & A'_{33} \end{pmatrix} \begin{pmatrix} x_{2,k}^+ \\ x_{3,k}^+ \end{pmatrix}.$$

As $\sigma(A'_{33}) \subset \mathcal{D}$ we have $x_{3,k}^+ \to 0$ as $k \to \infty$. From (4.6) we see that

$$J_T(x_0,u^+) = x_0^* P_{\mathcal{L}} x_0 - x_{3,T+1}^{+*}\Delta_{33}x_{3,T+1}^+ - x_{T+1}^{+*}(P_- - \lambda H^*H)x_{T+1}^+ - \lambda\|Hx_{T+1}^+\|^2.$$

As $J_T(x_0,u^+) - x_0^*P_{\mathcal{L}}x_0 \to 0$, $Hx_{T+1} \to 0$ and $x_{3,T+1}^{+*}\Delta_{33}x_{3,T+1}^+ \to 0$ as $T \to \infty$, we get $x_{T+1}^{+*}(P_- - \lambda H^*H)x_{T+1}^+ \to 0$. As $P_- - \lambda H^*H \le 0$ this gives $(P_- - \lambda H^*H)x_{T+1}^+ \to 0$ and hence $P_-x_T \to 0$. In turn this implies that $D_2x_{2,k}^+ + D_3x_{3,k}^+ \to 0$ (see (4.5)). As $x_{3,k} \to 0$ we obtain $D_2x_{2,k}^+ \to 0$. Using $x_{2,k+1}^+ = A_{22}x_{2,k}^+$ together with the fact that $\sigma(A_{22}) \subset \mathcal{T}$, this yields $D_2 = 0$ (note that x_0, and therefore $x_{2,0}$, is arbitrary). We conclude that

$$\ker \Delta = \mathcal{X}_2 \subseteq \ker \begin{pmatrix} P_- \\ H \end{pmatrix} = \mathcal{L} \cap \ker P_-.$$

Conversely, suppose that $\ker \Delta \subseteq \mathcal{L} \cap \ker P_-$. Then we have $P_{22}^- = 0$ and $P_{23}^- = 0$. Also $D_2 = 0$. Put $u = \{u_k\}$, where u_k is given by

$$u_k = -(R+B^*P_{\mathcal{L}}B)^{-1}(C+B^*P_{\mathcal{L}}A)x_k.$$

Then by (4.6)

$$J_T(x_0,u) = x_0^* P_{\mathcal{L}} x_0 - x_{3,T+1}^* P_{33}^+ x_{3,T+1}.$$

Since $x_{3,k+1} = A'_{33}x_{3,k}$ and $\sigma(A_{33})' \subset \mathcal{D}$, we have $x_{3,T+1} \to 0$. Hence $J_T(x_0, u) \to x_0^* P_{\mathcal{L}} x_0$, so $J(x_0, u) = x_0^* P_{\mathcal{L}} x_0$. Also note that

$$\begin{pmatrix} P_- \\ H \end{pmatrix} x_k = D_3 x_{3,k} \to 0$$

and hence $Hx_{3,k} \to 0$. Thus $u \in U_{\mathcal{L}}(x_0)$ and we can conclude that u is optimal.
The second part of (ii) was already proved (cf. (4.7)) □

REFERENCES

[1] W. A. Coppel, *Matrix quadratic equations*, Bull. Austral. Math. Soc., 10 (1974), pp. 377–401.
[2] P. Lancaster, L. Rodman, and A. C. M. Ran, *Hermitian solutions of the discrete algebraic Riccati equation*, Internat. J. Control, 44 (1986), pp. 777–802.
[3] A. C. M. Ran and L. Rodman, *Stability of invariant maximal semi-definite subspaces* I, Linear Algebra Appl., 62 (1984), pp. 51–86.
[4] ———, *Stable Hermitian solutions of discrete algebraic Riccati equation*, Math. Control Signals Systems, 5 (1992), pp. 165–193.
[5] H. L. Trentelman, *The regular free-endpoint linear quadratic problem with indefinite cost*, SIAM J. Control Optim., 27 (1989), pp. 27–42.
[6] J. M. Soethoudt and H. L. Trentelman, *The regular indefinite linear-quadratic problem with linear endpoint constraints*, Systems Control Lett., 12 (1989), pp. 23–31.
[7] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
[8] J. C. Willems, *Least squares stationary optimal control and the algebraic Riccati equation*, IEEE Trans. Automat. Control, 16 (1971), pp. 621–634.
[9] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, San Diego, CA, 1985.
[10] W. W. Lin, *A new method for computing the closed loop eigenvalues of a discrete time algebraic Riccati equation*, Linear Algebra Appl., 96 (1987), pp. 157–180.
[11] V. Mehrmann, *Existence, uniqueness and stability of solutions to singular linear quadratic optimal control problems*, Linear Algebra Appl., 121 (1989), pp. 291–321.
[12] T. Papas, A. J. Laub, and N. R. Sandell, *On the numerical solution of the discrete time algebraic Riccati equation*, IEEE Trans. Automat. Control, Vol. AC-25 (1980), pp. 631–641.